Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org
pp. **X–XX**

ON THE TRIALITY THEORY FOR A QUARTIC POLYNOMIAL OPTIMIZATION PROBLEM

David Yang Gao

School of Science, Information Technology and Engineering,
University of Ballarat, Victoria 3353, Australia.

Changzhi Wu

School of Science, Information Technology and Engineering,
University of Ballarat, Victoria 3353, Australia.
School of Mathematics, Chongqing Normal University,
Shapingba, Chongqing, 400047, China

(Communicated by Kok Lay Teo)

Abstract. This paper presents a detailed proof of the triality theorem for a class of fourth-order polynomial optimization problems. The method is based on linear algebra but it solves an open problem on the double-min duality left in 2003. Results show that the triality theory holds strongly in a tri-duality form if the primal problem and its canonical dual have the same dimension; otherwise, both the canonical min-max duality and the double-max duality still hold strongly, but the double-min duality holds weakly in a symmetrical form. Four numerical examples are presented to illustrate that this theory can be used to identify not only the global minimum, but also the largest local minimum and local maximum.

1. **Introduction and Motivation.** The concepts of *triality* and *tri-duality* were originally proposed in nonconvex mechanics [4, 5]. Mathematical theory of triality in its standard format is composed of three types of dualities: a canonical min-max duality and a pair of double-min and double-max dualities. The canonical min-max duality provides a sufficient condition for global minimum, while the double-min and double max dualities can be used to identify respectively the largest local minimum and local maximum. The tri-duality is a strong form of the triality principle [7]. Together with a *canonical dual transformation* and a *complementary-dual principle,* they comprise a versatile *canonical duality theory,* which can be used not only for solving a large class of challenging problems in nonconvex/nonsmooth analysis and continuous/discrete optimization [7, 8], but also for modeling complex systems and understanding multi-scale phenomena within a unified framework [4, 5, 7] (see also the review articles [10, 12, 17]).

For example, in the recent work by Gao and Ogden [13] on nonconvex variational/boundary value problems, it was discovered that both the global and local minimizers are usually nonsmooth functions and cannot be determined easily by traditional Newton-type numerical methods. However, by the canonical dual transformation, the nonlinear differential equation is equivalent to an algebraic equation, which can be solved analytically to obtain all solutions. Both global minimizer and local extrema were identified by the triality theory, which revealed some interesting phenomena in phase transitions.

The triality theory has attracted much attention recently in duality ways: Successful applications in multi-disciplinary fields of mathematics, engineering and sciences show that this theory is not only useful and versatile, but also beautiful in its mathematical format and rich in connotation

2000 *Mathematics Subject Classification.* 90C26, 90C30, 90C46, 90C49.

Key words and phrases. canonical duality, triality, global optimization, polynomial optimization, counter-examples.

of physics, which reveals a unified intrinsic duality pattern in complex systems; On the other hand, a large number of "counterexamples" have been presented in several papers since 2010. Unfortunately, most of these counterexamples are either fundamentally wrong (see [30, 31]), or repeatedly address an open problem left by Gao in 2003 on the double-min duality [9, 10].

The main goal of this paper is to solve this open problem left in 2003. The next section will present a brief review and the open problem in the triality theory. In Section 3, the triality theory is proved in its strong form as it was originally discovered. Section 4 shows that both the canonical min-max and the double-max dualities hold strongly in general, but the double-min duality holds weakly in a symmetrical form. Applications are illustrated in Section 5, where a linear perturbation method is used for solving certain critical problems. The paper ended by an Appendix and a section of concluding remarks.

2. **Canonical Duality Theory: A Brief Review and an Open Problem.** Let us begin with the general global extremum problem

$$(\mathcal{P}): \quad \text{ext}\left\{\Pi(\boldsymbol{x}) = \mathrm{W}(\boldsymbol{x}) + \frac{1}{2}\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{f}\rangle \mid \boldsymbol{x} \in \mathcal{X}_{\mathrm{a}}\right\}, \tag{1}$$

where $\mathcal{X}_a \subset \mathbb{R}^n$ is an open set, $\boldsymbol{x} = \{x_i\} \in \mathbb{R}^n$ is a decision vector, $\mathbf{A} = \{A_{ij}\} \in \mathbb{R}^{n\times n}$ is a given symmetric matrix, $\boldsymbol{f} = \{f_i\} \in \mathbb{R}^n$ is a given vector, and $\langle *, * \rangle$ denotes a bilinear form on $\mathbb{R}^n \times \mathbb{R}^n$; the function $W : \mathcal{X}_a \to \mathbb{R}$ is assumed to be nonconvex and differentiable (it is allowed to be nonsmooth and sub-differentiable for constrained problems). The notation ext$\{*\}$ stands for finding global extremal of the function given in $\{*\}$.

In this paper, we are interested only in three types of global extrema: the global minimum and a pair of the largest local minimum and local maximum. Therefore, the nonconvex term $W(\boldsymbol{x})$ in (1) is assumed to satisfy the objectivity condition [1], i.e., there exists a *(geometrically) nonlinear mapping* $\Lambda : \mathcal{X}_a \to \mathcal{V} \subset \mathbb{R}^m$ and a canonical function $V : \mathcal{V} \subset \mathbb{R}^m \to \mathbb{R}$ such that

$$W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x})) \quad \forall \boldsymbol{x} \in \mathcal{X}_a.$$

According to [7], a real valued function $V : \mathcal{V} \to \mathbb{R}$ is said to be a canonical function on its effective domain $\mathcal{V}_a \subset \mathcal{V}$ if its Legendre conjugate $V^* : \mathcal{V}^* \to \mathbb{R}$

$$V^*(\boldsymbol{\varsigma}) = \text{sta}\{\langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle - \mathrm{V}(\boldsymbol{\xi}) |\ \boldsymbol{\xi} \in \mathcal{V}_{\mathrm{a}}\} \tag{2}$$

is uniquely defined on its effective domain $\mathcal{V}_a^* \subset \mathcal{V}^*$ such that the canonical duality relations

$$\boldsymbol{\varsigma} = \nabla V(\boldsymbol{\xi}) \quad \Leftrightarrow \quad \boldsymbol{\xi} = \nabla V^*(\boldsymbol{\varsigma}) \quad \Leftrightarrow \quad V(\boldsymbol{\xi}) + V^*(\boldsymbol{\varsigma}) = \langle \boldsymbol{\xi}; \boldsymbol{\varsigma}\rangle \tag{3}$$

hold on $\mathcal{V}_a \times \mathcal{V}_a^*$, where $\langle *; * \rangle$ represents a bilinear form which puts $\mathcal{V}$ and $\mathcal{V}^*$ in duality. The notation sta$\{*\}$ stands for solving the stationary point problem in $\{*\}$. By this one-to-one canonical duality, the nonconvex function $W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x}))$ can be replaced by $\langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma}\rangle - V^*(\boldsymbol{\varsigma})$ such that the nonconvex function $\Pi(\boldsymbol{x})$ in (1) can be written as

$$\Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) = \langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma}\rangle - V^*(\boldsymbol{\varsigma}) + \frac{1}{2}\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{f}\rangle, \tag{4}$$

which is the so-called *total complementary (energy) function* introduced by Gao and Strang in 1989. By using this total complementary function, the canonical dual function $\Pi^d : \mathcal{V}_a^* \to \mathbb{R}$ can be formulated as

$$\Pi^d(\boldsymbol{\varsigma}) = \text{sta}\{\Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) |\ \forall \boldsymbol{x} \in \mathcal{X}_{\mathrm{a}}\} = \mathrm{U}^\Lambda(\boldsymbol{\varsigma}) - \mathrm{V}^*(\boldsymbol{\varsigma}), \tag{5}$$

where $U^\Lambda : \mathcal{V}_a^* \to \mathbb{R}$ is called the Λ-conjugate of $U(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f}\rangle - \frac{1}{2}\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle$, defined by [7] as

$$U^\Lambda(\boldsymbol{\varsigma}) = \text{sta}\{\langle \Lambda(\boldsymbol{x}); \boldsymbol{\varsigma}\rangle - \mathrm{U}(\boldsymbol{x}) |\ \boldsymbol{x} \in \mathcal{X}_{\mathrm{a}}\}. \tag{6}$$

Let $\mathcal{S}_a \subset \mathcal{V}_a^*$ be the feasible domain of $U^\Lambda(\boldsymbol{\varsigma})$; then the canonical dual problem is to solve the stationary point problem

$$(\mathcal{P}^d): \quad \text{ext}\{\ \Pi^{\mathrm{d}}(\boldsymbol{\varsigma}) |\ \boldsymbol{\varsigma} \in \mathcal{S}_{\mathrm{a}}\}. \tag{7}$$

---

[1]The concept of objectivity in science means that qualitative and quantitative descriptions of physical phenomena remain unchanged when the phenomena are observed under a variety of conditions. That is, the objective function should be independent with the choice of the coordinate systems. In continuum mechanics, the objectivity is also regarded as the *principle of frame-indifference.* See Chapter 6 in [7] for mathematical definitions of the objectivity and geometric nonlinearity in differential geometry and finite deformation field theory. Detailed discussion of objectivity in global optimization will be given in another paper [18].

Theorem 2.1 (Complementary-duality principle [7]). *Problem $(\mathcal{P}^d)$ is a canonical dual to $(\mathcal{P})$ in the sense that if $(\bar{\boldsymbol{x}}, \bar{\varsigma})$ is a critical point of $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$, then $\bar{\boldsymbol{x}}$ is a critical point of $(\mathcal{P})$, $\bar{\boldsymbol{\varsigma}}$ is a critical point of $(\mathcal{P}^d)$, and*

$$\Pi(\bar{\boldsymbol{x}}) = \Xi(\bar{\boldsymbol{x}}, \bar{\varsigma}) = \Pi^d(\bar{\varsigma}). \tag{8}$$

Theorem 2.1 implies a perfect duality relation (i.e. no duality gap) between the primal problem and its canonical dual[2]. The formulation of $\Pi^d(\boldsymbol{\varsigma})$ depends on the geometrical operator $\Lambda(\boldsymbol{x})$. In many applications, the geometrical operator Λ is usually a quadratic mapping over a given field [7]. In finite dimensional space, this quadratic operator can be written as a vector-valued function (see [8], page 150)

$$\Lambda(\boldsymbol{x}) = \left\{ \frac{1}{2}\boldsymbol{x}^T \mathbf{B}^k \boldsymbol{x} \right\}_{k=1}^m : \ \mathcal{X}_a \subset \mathbb{R}^n \to \mathcal{V}_a \subset \mathbb{R}^m, \tag{9}$$

where $\mathbf{B}^k = \{\boldsymbol{B}_{ij}^k\} \in \mathbb{R}^{n\times n}$ is a symmetrical matrix for each $k = 1, 2, \cdots, m$, and $\mathcal{V}_a \subset \mathbb{R}^m$ is defined by

$$\mathcal{V}_a = \left\{ \boldsymbol{\xi} \in \mathbb{R}^m | \ \xi_k = \frac{1}{2}\boldsymbol{x}^T \mathbf{B}^k \boldsymbol{x} \ \ \forall \boldsymbol{x} \in \mathcal{X}_a, \ \ k = 1, \dots, m \right\}.$$

In this case, the total complementary function has the form

$$\Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) = \frac{1}{2}\langle \boldsymbol{x}, \mathbf{G}(\boldsymbol{\varsigma})\boldsymbol{x}\rangle - V^*(\boldsymbol{\varsigma}) - \langle \boldsymbol{x}, \boldsymbol{f}\rangle, \tag{10}$$

where $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^{n\times n}$ is a matrix-valued function defined by

$$\mathbf{G}(\boldsymbol{\varsigma}) = \mathbf{A} + \sum_{k=1}^m \varsigma_k \mathbf{B}^k. \tag{11}$$

The critical condition $\nabla_{\boldsymbol{x}} \Xi(\boldsymbol{x}, \boldsymbol{\varsigma}) = 0$ leads to the canonical equilibrium equation

$$\mathbf{G}(\boldsymbol{\varsigma})\boldsymbol{x} = \boldsymbol{f}. \tag{12}$$

Clearly, for any given $\boldsymbol{\varsigma} \in \mathcal{V}_a^*$, if the vector $\boldsymbol{f} \in \mathcal{C}_{ol}(\mathbf{G}(\boldsymbol{\varsigma}))$, where $\mathcal{C}_{ol}(\mathbf{G})$ stands for a space spanned by the columns of $\mathbf{G}$, the canonical equilibrium equation (12) can be solved analytically as[3] $\boldsymbol{x} = [\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\boldsymbol{f}$. Therefore, the canonical dual feasible space $\mathcal{S}_a \subset \mathcal{V}_a^*$ can be defined as

$$\mathcal{S}_a = \{\boldsymbol{\varsigma} \in \mathcal{V}_a^* \mid \boldsymbol{f} \in \mathcal{C}_{ol}(\mathbf{G}(\boldsymbol{\varsigma}))\},$$

and on $\mathcal{S}_a$ the canonical dual $\Pi^d(\boldsymbol{\varsigma})$ is well-defined as

$$\Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\langle \mathbf{G}(\boldsymbol{\varsigma})^{-1}\boldsymbol{f}, \boldsymbol{f}\rangle - V^*(\boldsymbol{\varsigma}). \tag{13}$$

Theorem 2.2 (Analytic solution [7]). *If $\bar{\varsigma} \in \mathcal{S}_a$ is a critical solution of $(\mathcal{P}^d)$, then*

$$\bar{\boldsymbol{x}} = \mathbf{G}(\bar{\varsigma})^{-1}\boldsymbol{f} \tag{14}$$

is a critical solution of $(\mathcal{P})$ and $\Pi(\bar{\boldsymbol{x}}) = \Pi^d(\bar{\varsigma})$.

Conversely, if $\bar{\boldsymbol{x}}$ is a critical solution of $(\mathcal{P})$, it must be in the form of (14) for a certain critical solution $\bar{\varsigma}$ of $(\mathcal{P}^d)$.

The canonical dual function $\Pi^d(\boldsymbol{\varsigma})$ for a general quadratic operator Λ was first formulated in nonconvex analysis, where Theorem 2.2 is called the pure complementary energy principle, [6]. In finite deformation theory, this theorem solved an open problem left by Hellinger (1914) and Reissner (1954) (see [25]). The analytical solution theorem has been successfully applied for solving a class of nonconvex problems in mathematical physics, including Einstein's special relativity theory [7], nonconvex mechanics and phase transitions in solids [13]. In global optimization, the primal solutions to nonconvex minimization and integer programming problems are usually located on the boundary of the feasible space. By Theorem 2.2, these solutions can be analytically determined by critical points of the canonical dual function $\Pi^d(\boldsymbol{\varsigma})$ (see [3, 11, 14, 16]).

---

[2]The complementary-dual in physics means perfect dual in optimization, i.e., the canonical dual in Gao's work, which means no duality gap. Otherwise, any duality gap will violet the energy conservation law. Therefore, each complementary-dual variational statement in continuum mechanics is usually refereed as a principle.

[3]In this paper $\mathbf{G}^{-1}$ should be understood as the generalized inverse if $\det \mathbf{G} = 0$ [8].

In order to identify both global and local extrema of the primal and dual problems, we assume, without losing much generality, that the canonical function $V : \mathcal{E}_a \to \mathbb{R}$ is convex and let

$$\mathcal{S}_a^+ = \{\varsigma \in \mathcal{S}_a \mid \mathbf{G}(\varsigma) \succeq 0\}, \tag{15}$$

$$\mathcal{S}_a^- = \{\varsigma \in \mathcal{S}_a \mid \mathbf{G}(\varsigma) \prec 0\}, \tag{16}$$

where $\mathbf{G}(\varsigma) \succeq 0$ means that $\mathbf{G}(\varsigma)$ is a positive semi-definite matrix and $\mathbf{G}(\varsigma) \prec 0$ means that $\mathbf{G}(\varsigma)$ is negative definite.

Theorem 2.3 (Triality Theorem [8]). *Let $(\bar{\boldsymbol{x}}, \bar{\varsigma})$ be a critical point of $\Xi(\boldsymbol{x}, \varsigma)$.*

If $\mathbf{G}(\bar{\varsigma}) \succeq 0$, then $\bar{\varsigma}$ is a global maximizer of Problem $(\mathcal{P}^d)$, the vector $\bar{\boldsymbol{x}}$ is a global minimizer of Problem $(\mathcal{P})$, and the following canonical min-max duality statement holds:

$$\min_{\boldsymbol{x} \in \mathcal{X}_a} \Pi(\boldsymbol{x}) = \Xi(\bar{\boldsymbol{x}}, \bar{\varsigma}) = \max_{\varsigma \in \mathcal{S}_a^+} \Pi^d(\varsigma). \tag{17}$$

If $\mathbf{G}(\bar{\varsigma}) \prec 0$, then there exists a neighborhood $\mathcal{X}_o \times \mathcal{S}_o \subset \mathcal{X}_a \times \mathcal{S}_a^-$ of $(\bar{\boldsymbol{x}}, \bar{\varsigma})$ for which we have either the double-min duality statement

$$\min_{\boldsymbol{x} \in \mathcal{X}_o} \Pi(\boldsymbol{x}) = \Xi(\bar{\boldsymbol{x}}, \bar{\varsigma}) = \min_{\varsigma \in \mathcal{S}_o} \Pi^d(\varsigma), \tag{18}$$

or the double-max duality statement

$$\max_{\boldsymbol{x} \in \mathcal{X}_o} \Pi(\boldsymbol{x}) = \Xi(\bar{\boldsymbol{x}}, \bar{\varsigma}) = \max_{\varsigma \in \mathcal{S}_o} \Pi^d(\varsigma). \tag{19}$$

The triality theory provides actual global extremum criteria for three types of solutions to the nonconvex problem $(\mathcal{P})$: a global minimizer $\bar{\boldsymbol{x}}(\bar{\varsigma})$ if $\bar{\varsigma} \in \mathcal{S}_a^+$ and a pair of the largest-valued local extrema. In other words, $\bar{\boldsymbol{x}}(\bar{\varsigma})$ is the largest-valued local maximizer if $\bar{\varsigma} \in \mathcal{S}_a^-$ is a local maximizer; $\bar{\boldsymbol{x}}(\bar{\varsigma})$ is the largest-valued local minimizer if $\bar{\varsigma} \in \mathcal{S}_a^-$ is a local minimizer. This pair of largest local extrema plays a critical role in nonconvex analysis of post-bifurcation and phase transitions.

Remark 1 (Relation between Lagrangian Duality and Canonical Duality).

The main difference between the Lagrangian-type dualities (including the equivalent Fenchel-Moreau-Rockfellar dualities) and the canonical duality is the operator $\Lambda : \mathcal{X}_a \to \mathcal{V}_a$. In fact, if Λ is linear, the primal problem $(\mathcal{P})$ is called geometrically linear in [7] *and the total complementary function $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ is simply the well-known Lagrangian and is denoted as*

$$\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\varsigma}) = \langle \Lambda \boldsymbol{x}; \boldsymbol{\varsigma} \rangle - V^*(\boldsymbol{\varsigma}) - F(\boldsymbol{x}). \tag{20}$$

In convex (static) systems, $F(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f} \rangle$ is linear and $\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\varsigma})$ is a saddle function. Therefore, the well-known saddle min-max duality links a convex minimization problem $(\mathcal{P})$ to a concave maximization dual problem with linear constraint:

$$\max \{\Pi^*(\boldsymbol{\varsigma}) = -V^*(\boldsymbol{\varsigma}) | \ \Lambda^* \boldsymbol{\varsigma} = \boldsymbol{f}, \ \ \boldsymbol{\varsigma} \in \mathcal{V}_a^*\}, \tag{21}$$

where Λ^* is the conjugate operator of Λ defined via $\langle \Lambda \boldsymbol{x}; \boldsymbol{\varsigma} \rangle = \langle \boldsymbol{x}, \Lambda^* \boldsymbol{\varsigma} \rangle$. Using the Lagrange multiplier $\boldsymbol{x} \in \mathcal{X}_a$ to relax the equality constraint, the Lagrangian $\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\varsigma})$ is obtained. By the fact that the (canonical) duality in convex static systems is unique, the saddle min-max duality is refereed as the mono-duality *in complex systems (see Chapter 1 in* [7]*).*

Since the linear operator Λ can not change the convexity of $W(\boldsymbol{x}) = V(\Lambda \boldsymbol{x})$, the Lagrangian duality theory can be used mainly for convex problems. It is known that if $W(\boldsymbol{x})$ is nonconvex, then the Lagrangian duality as well as the related Fenchel-Moreau-Rockafellar duality will produce the so-called duality gap. Comparing the canonical dual function $\Pi^d(\boldsymbol{\varsigma})$ in (13) with the Lagrangian dual function $\Pi^*(\boldsymbol{\varsigma})$ in (21), we know that the duality gap is $\frac{1}{2}\langle \mathbf{G}(\boldsymbol{\varsigma})^{-1} \boldsymbol{f}, \boldsymbol{f} \rangle$.

The canonical duality theory is based on the (geometrically) nonlinear mapping $\Lambda : \mathcal{X}_a \to \mathcal{V}_a$ and the canonical transformation $W(\boldsymbol{x}) = V(\Lambda(\boldsymbol{x}))$. The total complementary function $\Xi(\boldsymbol{x}, \boldsymbol{\varsigma})$ is also known as the nonlinear or extended Lagrangian and is denoted by $L(\boldsymbol{x}, \boldsymbol{\varsigma})$ due to the geometric nonlinearity of $\Lambda(\boldsymbol{x})$ (see [7, 10]*). Relations between the canonical duality and the classical Lagrangian duality are discussed in* [16].

Remark 2 (Geometrical Nonlinearity and Complementary Gap Function). *The canonical min-max duality statement (17) was first proposed by Gao and Strang in nonconvex/nonsmooth analysis and mechanics in 1989* [19]*, where $\Pi(\boldsymbol{x}) = W(\boldsymbol{x}) - F(\boldsymbol{x})$ is the so-called total potential energy with W representing the internal (or stored) energy and F the external energy. The* geometrical nonlinearity *is a standard terminology in finite deformation theory, which implies that the geometrical equation (or the configuration-strain relation) $\boldsymbol{\xi} = \Lambda(\boldsymbol{x})$ is nonlinear. By definition in physics, a function $F(\boldsymbol{x})$ is called the external energy means that its (sub-)differential must be the*

external force (or input) $\boldsymbol{f}$. Therefore, in Gao and Strang's work, the external energy should be a linear function(al) $F(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f} \rangle$ on its effective domain. In this case, the matrix $\mathbf{G}(\boldsymbol{\varsigma})$ is a Hessian of the so-called complementary gap function *(i.e. the Gao-Strang gap function [19])*

$$G_{ap}(\boldsymbol{x}, \boldsymbol{\varsigma}) = \langle -\Lambda_c(\boldsymbol{x}); \boldsymbol{\varsigma} \rangle, \tag{22}$$

where $\Lambda_c(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{x}^T\mathbf{B}^k\boldsymbol{x}$ is called the complementary operator of a Gâteauxdifferential $\Lambda_t(\boldsymbol{x})\boldsymbol{x} = \boldsymbol{x}^T\mathbf{B}^k\boldsymbol{x}$ of $\Lambda(\boldsymbol{x})$ [19]. Actually, in Gao and Strang's original work, the canonical min-max duality statement holds in a general (weak) condition, i.e., $G_{ap}(\boldsymbol{x}, \bar{\boldsymbol{\varsigma}}) \geq 0$, $\forall \boldsymbol{x} \in \mathcal{X}_a$ in field theory (corresponding to the strong condition $\mathbf{G}(\bar{\boldsymbol{\varsigma}}) \succeq 0$). The related canonical duality theory has been generalized to nonconvex variational analysis of a large deformation (von Karman) plate (where $\mathbf{A} = \Delta^2$ [34]), nonconvex (chaotic) dynamical systems (where $\mathbf{A} = \Delta - \partial^2/\partial t^2$ [10]), and general nonconvex constrained problems in global optimization. Since $F(\boldsymbol{x})$ in these general applications is the quadratic function $-\frac{1}{2}\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x} \rangle + \langle \boldsymbol{x}, \boldsymbol{f} \rangle$, the Gao-Strang gap function (22) should be replaced by the generalized form $G_{ap}(\boldsymbol{x}, \boldsymbol{\varsigma}) = \frac{1}{2}\langle \boldsymbol{x}, \mathbf{G}(\boldsymbol{\varsigma})\boldsymbol{x} \rangle$ (see the review article by Gao and Sherali [17]). This gap function recovers the existing duality gap in traditional duality theories and provides a sufficient global optimality condition for general nonconvex problems in both infinite and finite dimensional systems (see review articles [10, 17]). By the fact that the geometrical mapping Λ in Gao and Strang's work is a tensor-like operator, it has been realized recently that the popular semi-definite programming method is actually a special application (where $W(\boldsymbol{x})$ is a quadratic function) of the canonical min-max duality theory proposed in 1989 (see [14, 15]).

In a recent paper by Voisei and Zalinescu [31], they unfortunately misunderstood some basic terminologies in continuum physics, such as geometric nonlinearity, internal and external energies, and present "counterexamples" to the Gao-Strang theory based on certain "artificially chosen" operators $\Lambda(\boldsymbol{x})$ and quadratic functions $F(\boldsymbol{x})$. Whereas in the stated contexts, the geometrical operator Λ should be a canonical measure (Cauchy-Reimann type finite deformation operator, see Chapter 6 in [7]) and the external energy $F(\boldsymbol{x})$ is typically a linear functional on its effective domain; otherwise, its (sub)-differential will not be the external force. Interested readers are refereed by [18] for further discussion.

Remark 3 (Double-Min Duality and Open Problem)**.** *The double-min and double-max duality statements were discovered simultaneously in a post-buckling analysis of large deformed beam model [4, 5] in 1996, where the finite strain measure Λ is a quadratic differential operator from a 2-D displacement field to a 2-D canonical strain field. Therefore, the triality theory was first proposed in its strong form, i.e. the so-called* tri-duality theory *(see the next section). Later on when Gao was writing his duality book [7], he realized that this pair of double-min and double-max dualities holds naturally in convex Hamilton systems. Accordingly, a* bi-duality theorem *was proposed and proved for geometrically linear systems (where Λ is a linear operator; see Chapter 2 in [7]). Following this, the triality theory was naturally generalized to geometrically nonlinear systems (nonlinear Λ; see Chapter 3 in [7]) with applications to global optimization problems [8]. However, it was discovered in 2003 that if $n \neq m$ in the quadratic mapping (9), the double-min duality statement needs "certain additional constraints". For the sake of mathematical rigor, the double-min duality was not included in the triality theory and these additional constraints were left as an open problem (see Remark 1 in [9], also Theorem 3 and its Remark in a review article by Gao [10]). By the fact that the double-max duality is always true, the double-min duality was still included in the triality theory in the "either-or" form in many applications (see [12, 16]). However, ignoring the open problem related to the "certain additional constraints" on the double-min duality statement has led to some misleading results.*

The goal of this paper is to solve this open problem by providing a simple proof of the triality theory based on linear algebra. To help understanding the intrinsic characteristics of the original problem and its canonical dual, we assume that the nonconvex objective function $W(\boldsymbol{x})$ is a sum of fourth-order canonical polynomials

$$W(\boldsymbol{x}) = \frac{1}{2}\sum_{k=1}^{m} \beta_k \left(\frac{1}{2}\boldsymbol{x}^T\mathbf{B}^k\boldsymbol{x} - d^k\right)^2, \tag{23}$$

where $\mathbf{B}^k = \left\{B^k_{ij}\right\} \in \mathbb{R}^{n\times n}$, $k = 1, \cdots, m$, are all symmetric matrices, $\beta_k > 0$ and $d^k \in \mathbb{R}$, $k = 1, \cdots, m$ are given constants. This polynomial is actually a discretized form of the so-called *double-well potential*, first proposed by van der Waals in thermodynamics in 1895 (see [26]), which is the mathematical model for natural phenomena of bifurcation and phase transitions in biology,

chemistry, cosmology, continuum mechanics, material science, and quantum field theory, etc. (see [5, 20, 23, 24]). By using the quadratic geometrical operator $\Lambda(\boldsymbol{x})$ given by (9), the canonical function

$$V(\boldsymbol{\xi}) = \frac{1}{2}(\boldsymbol{\xi}-\boldsymbol{d})^T\boldsymbol{\beta}(\boldsymbol{\xi}-\boldsymbol{d}) \tag{24}$$

and its Legendre conjugate

$$V^*(\boldsymbol{\varsigma}) = \frac{1}{2}\boldsymbol{\varsigma}^T\boldsymbol{\beta}^{-1}\boldsymbol{\varsigma} + \boldsymbol{\varsigma}^T\boldsymbol{d} \tag{25}$$

are quadratic functions, where $\boldsymbol{\beta} = \text{Diag}\,(\beta^k)$ represents the diagonal matrix defined by the non-zero vector $\{\beta^k\}$.

In the following discussions, we assume that all the critical points of problem $(\mathcal{P})$ are non-singular, i.e., if $\nabla\Pi(\bar{\boldsymbol{x}}) = 0$, then

$$\det\nabla^2\Pi(\bar{\boldsymbol{x}}) \neq 0. \tag{26}$$

We will first prove that if $n = m$, the triality theorem holds in its strong form; otherwise, the theorem holds in its weak form. Three numerical examples are used to illustrate the effectiveness and efficiency of the canonical duality theory.

3. **Strong Triality Theory for Quartic Polynomial Optimization: Tri-Duality Theorem.** We first consider the case $m = n$. For simplicity, we assume that $\beta_k = 1$ in the following discussion (otherwise, $\mathbf{B}^k$ can be replaced by $\sqrt{\beta_k}\mathbf{B}^k$ and d^k is replaced by $d^k/\sqrt{\beta_k}$). In this case, the problem (1) is denoted as problem $(\mathcal{P})$. Its canonical dual is

$$(\mathcal{P}^d): \quad \text{ext} \quad \left\{\Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\boldsymbol{f}^T[\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\boldsymbol{f} - \frac{1}{2}\boldsymbol{\varsigma}^T\boldsymbol{\varsigma} - \boldsymbol{\varsigma}^T\boldsymbol{d} \mid \boldsymbol{\varsigma}\in\mathcal{S}_a\subset\mathbb{R}^n\right\}. \tag{27}$$

Theorem 3.1 (Tri-Duality Theorem)**.**

Suppose that $m = n$, that the assumption (26) is satisfied, that $\bar{\boldsymbol{\varsigma}}$ is a critical point of Problem $(\mathcal{P}^d)$ and that $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f}$.

If $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^+$, then $\bar{\boldsymbol{\varsigma}}$ is a global maximizer of Problem $(\mathcal{P}^d)$ in $\mathcal{S}_a^+$ if and only if $\bar{\boldsymbol{x}}$ is a global minimizer of Problem $(\mathcal{P})$, i.e., the following canonical min-max statement holds:

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathbb{R}^n}\Pi(\boldsymbol{x}) \Longleftrightarrow \max_{\boldsymbol{\varsigma}\in\mathcal{S}_a^+}\Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{28}$$

On the other hand, if $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$, then, there exists a neighborhood $\mathcal{X}_o\times\mathcal{S}_o\subset\mathbb{R}^n\times\mathcal{S}_a^-$ of $(\bar{\boldsymbol{x}},\bar{\boldsymbol{\varsigma}})$, such that either one of the following two statements holds.

(A) The double-min duality statement

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathcal{X}_o}\Pi(\boldsymbol{x}) \iff \min_{\boldsymbol{\varsigma}\in\mathcal{S}_o}\Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}), \tag{29}$$

or (B) the double-max duality statement

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathcal{X}_o}\Pi(\boldsymbol{x}) \iff \max_{\boldsymbol{\varsigma}\in\mathcal{S}_o}\Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{30}$$

Proof. If $\bar{\boldsymbol{\varsigma}}$ is a critical point of the canonical dual problem $(\mathcal{P}^d)$, the criticality condition

$$\nabla\Pi^d(\boldsymbol{\varsigma}) = \frac{1}{2}\begin{bmatrix} \boldsymbol{f}^T[\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\mathbf{B}^1[\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\boldsymbol{f} \\ \cdots \\ \boldsymbol{f}^T[\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\mathbf{B}^n[\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\boldsymbol{f}\end{bmatrix} - \boldsymbol{\varsigma} - \boldsymbol{d} = 0 \in\mathbb{R}^n \tag{31}$$

leads to $\bar{\boldsymbol{\varsigma}} = \Lambda(\bar{\boldsymbol{x}})$. By the fact that $\nabla\Pi(\bar{\boldsymbol{x}}) = \mathbf{G}(\bar{\boldsymbol{\varsigma}})\bar{\boldsymbol{x}} - \boldsymbol{f} = 0\in\mathbb{R}^n$, it follows that $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f}$ is a critical point of Problem $(\mathcal{P})$.

To prove the validity of the canonical min-max statement (28), let $\bar{\boldsymbol{\varsigma}}$ be a critical point and $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^+$. Since $\Pi^d(\boldsymbol{\varsigma})$ is concave on $\mathcal{S}_a^+$, the critical point $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^+$ must be a global maximizer of $\Pi^d(\boldsymbol{\varsigma})$ on $\mathcal{S}_a^+$.

On the other hand, by the convexity of $V(\boldsymbol{\xi})$, we have

$$V(\boldsymbol{\xi}) - V(\bar{\boldsymbol{\xi}}) \geq \langle\boldsymbol{\xi}-\bar{\boldsymbol{\xi}};\nabla V(\bar{\boldsymbol{\xi}})\rangle = \langle\boldsymbol{\xi}-\bar{\boldsymbol{\xi}};\bar{\boldsymbol{\varsigma}}\rangle. \tag{32}$$

Substituting $\boldsymbol{\xi} = \Lambda(\boldsymbol{x})$ and $\bar{\boldsymbol{\xi}} = \Lambda(\bar{\boldsymbol{x}})$ into (32), we obtain

$$V(\Lambda(\boldsymbol{x})) - V(\Lambda(\bar{\boldsymbol{x}})) \geq \langle\Lambda(\boldsymbol{x}) - \Lambda(\bar{\boldsymbol{x}});\bar{\boldsymbol{\varsigma}}\rangle.$$

This leads to

$$\Pi(\boldsymbol{x}) - \Pi(\bar{\boldsymbol{x}}) \geq \langle\Lambda(\boldsymbol{x}) - \Lambda(\bar{\boldsymbol{x}});\bar{\boldsymbol{\varsigma}}\rangle + \frac{1}{2}\langle\boldsymbol{x},\mathbf{A}\boldsymbol{x}\rangle - \frac{1}{2}\langle\bar{\boldsymbol{x}},\mathbf{A}\bar{\boldsymbol{x}}\rangle - \langle\boldsymbol{x}-\bar{\boldsymbol{x}},\boldsymbol{f}\rangle,\ \forall\boldsymbol{x}\in\mathbb{R}^n. \tag{33}$$

By the fact that

$$\bar{\boldsymbol{\varsigma}} = \Lambda(\bar{\boldsymbol{x}}) - \boldsymbol{d}, \tag{34}$$

we have

$$\Pi(\boldsymbol{x}) - \Pi(\bar{\boldsymbol{x}}) \geq \frac{1}{2}\langle \boldsymbol{x}, \mathbf{G}(\bar{\varsigma})\boldsymbol{x}\rangle - \frac{1}{2}\langle \bar{\boldsymbol{x}}, \mathbf{G}(\bar{\varsigma})\bar{\boldsymbol{x}}\rangle - \langle \boldsymbol{x}-\bar{\boldsymbol{x}}, \mathbf{G}(\bar{\varsigma})\bar{\boldsymbol{x}}\rangle. \tag{35}$$

For a fixed $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^+$, the convexity of the complementary gap function $\boldsymbol{G}_{ap}(\boldsymbol{x}, \bar{\boldsymbol{\varsigma}}) = \frac{1}{2}\langle \boldsymbol{x}, \mathbf{G}(\bar{\varsigma})\boldsymbol{x}\rangle$ on $\mathcal{X}_a$ leads to

$$\boldsymbol{G}_{ap}(\boldsymbol{x}, \bar{\boldsymbol{\varsigma}}) - \boldsymbol{G}_{ap}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}}) \geq \langle \boldsymbol{x}-\bar{\boldsymbol{x}}, \nabla_{\boldsymbol{x}}\boldsymbol{G}_{ap}(\bar{\boldsymbol{x}}, \bar{\varsigma})\rangle = \langle \boldsymbol{x}-\bar{\boldsymbol{x}}, \mathbf{G}(\bar{\varsigma})\bar{\boldsymbol{x}}\rangle \;\; \forall \boldsymbol{x} \in \mathbb{R}^n. \tag{36}$$

Therefore, we have

$$\Pi(\boldsymbol{x}) - \Pi(\bar{\boldsymbol{x}}) \geq \langle \boldsymbol{x}-\bar{\boldsymbol{x}}, \mathbf{G}(\bar{\varsigma})\bar{\boldsymbol{x}}\rangle - \langle \boldsymbol{x}-\bar{\boldsymbol{x}}, \mathbf{G}(\bar{\varsigma})\bar{\boldsymbol{x}}\rangle = 0 \;\; \forall \boldsymbol{x} \in \mathbb{R}^n. \tag{37}$$

This shows that $\bar{\boldsymbol{x}}$ is a global minimizer of Problem $(\mathcal{P})$. Since it is assumed that $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^+$, it follows that (28) is satisfied.

We move on to prove the double-min duality statement (29).

Let $\bar{\boldsymbol{\varsigma}}$ be a critical point of $\Pi^d(\boldsymbol{\varsigma})$ and $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$. It is easy to verify that

$$\nabla\Pi(\boldsymbol{x}) = \sum_{k=1}^{n}\left(\frac{1}{2}\boldsymbol{x}^T\mathbf{B}^k\boldsymbol{x} - d^k\right)\mathbf{B}^k\boldsymbol{x} + \mathbf{A}\boldsymbol{x} - \boldsymbol{f},$$

$$\nabla^2\Pi(\bar{\boldsymbol{x}}) = \mathbf{G}(\bar{\varsigma}) + \mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T, \tag{38}$$

where

$$\mathbf{F}(\boldsymbol{x}) = \left[\mathbf{B}^1\boldsymbol{x}, \mathbf{B}^2\boldsymbol{x}, \cdots, \mathbf{B}^n\boldsymbol{x}\right].$$

In light of (31), $\nabla^2\Pi^d(\bar{\varsigma})$ can be expressed in terms of $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$ as follows:

$$\nabla^2\Pi^d(\bar{\varsigma}) = -\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}}) - \mathbf{I},$$

where $\mathbf{I}$ is the identity matrix. If the critical point $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ is a local minimizer, we have $\nabla^2\Pi^d(\bar{\varsigma}) \succeq 0$. This leads to

$$-\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}}) \succeq \mathbf{I}. \tag{39}$$

Therefore, $-\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}})$ is positive definite and $\boldsymbol{F}(\bar{\boldsymbol{x}})$ is invertible. By multiplying $\left(\mathbf{F}(\bar{\boldsymbol{x}})^T\right)^{-1}$ and $\mathbf{F}(\bar{\boldsymbol{x}})^{-1}$ to the left and right sides of (39), respectively, we obtain

$$-[\mathbf{G}(\bar{\varsigma})]^{-1} \succeq \left(\mathbf{F}(\bar{\boldsymbol{x}})^T\right)^{-1}(\mathbf{F}(\bar{\boldsymbol{x}}))^{-1}.$$

According to Lemma 6.2 in Appendix, the following matrix inequality is obtained:

$$\nabla^2\Pi(\bar{\boldsymbol{x}}) = \mathbf{G}(\bar{\varsigma}) + \mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T \succeq 0.$$

By the assumption (26), $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$ is also a local minimizer of Problem $(\mathcal{P})$. Therefore, on a neighborhood $\mathcal{X}_o \times \mathcal{S}_o \subset \mathbb{R}^n \times \mathcal{S}_a^-$ of $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$, we have

$$\min_{\boldsymbol{x}\in\mathcal{X}_o}\Pi(\boldsymbol{x}) = \Xi(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}}) = \min_{\boldsymbol{\varsigma}\in\mathcal{S}_o}\Pi^d(\boldsymbol{\varsigma}).$$

Similarly, we can show that if $\bar{\boldsymbol{x}}$ is a local minimizer of Problem $(\mathcal{P})$, the corresponding $\bar{\boldsymbol{\varsigma}}$ is also a local minimizer of Problem $(\mathcal{P}^d)$.

The next task is to prove the double-max duality statement (30).

Let $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ be a local maximizer of Problem $(\mathcal{P}^d)$. Then, we have $\nabla^2\Pi^d(\bar{\varsigma}) \preceq 0$. This gives us

$$\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}}) + \mathbf{I} \succeq 0. \tag{40}$$

Now we have two possible cases regarding the invertibility of $\mathbf{F}(\bar{\boldsymbol{x}})$. If $\mathbf{F}(\bar{\boldsymbol{x}})$ is invertible, then by using a similar argument as presented above, we can show that the relations

$$\max_{\boldsymbol{x}\in\mathcal{X}_o}\Pi(\boldsymbol{x}) = \Xi(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}}) = \max_{\boldsymbol{\varsigma}\in\mathcal{S}_o}\Pi^d(\boldsymbol{\varsigma})$$

hold on a neighborhood $\mathcal{X}_o \times \mathcal{S}_o \subset \mathbb{R}^n \times \mathcal{S}_a^-$ of $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{\varsigma}})$.

If $\mathbf{F}(\bar{\boldsymbol{x}})$ is not invertible, by Lemma 6.1 in the Appendix, there exists two orthogonal matrices $\mathbf{E}$ and $\mathbf{K}$ such that

$$\mathbf{F}(\bar{\boldsymbol{x}}) = \mathbf{EDK}, \tag{41}$$

where $\mathbf{E}^T\mathbf{E} = \mathbf{I} = \mathbf{K}^T\mathbf{K}$ and $\mathbf{D} = \mathrm{Diag}(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and $r = \mathrm{rank}\,(\mathbf{F}(\bar{\boldsymbol{x}}))$. Substituting (41) into (40), we obtain

$$-\mathbf{K}^T\mathbf{D}^T\mathbf{E}^T[\mathbf{G}(\varsigma)]^{-1}\mathbf{EDK} - \mathbf{I} \preceq 0. \tag{42}$$

Thus,

$$-\mathbf{D}^T\left[\mathbf{E}^T\mathbf{G}(\bar{\varsigma})\mathbf{E}\right]^{-1}\mathbf{D}-\mathbf{I}\preceq 0. \tag{43}$$

Applying Lemma 6.4 in Appendix to (43), it follows that

$$\mathbf{E}^T\mathbf{G}(\bar{\varsigma})\mathbf{E}+\mathbf{D}\mathbf{D}^T=\mathbf{E}^T\mathbf{G}(\bar{\varsigma})\mathbf{E}+\mathbf{D}\mathbf{K}\mathbf{K}^T\mathbf{D}^T\preceq 0.$$

Finally, we have

$$\nabla^2\Pi(\bar{\boldsymbol{x}})=\mathbf{G}(\bar{\varsigma})+\mathbf{E}\mathbf{D}\mathbf{K}\mathbf{K}^T\mathbf{D}^T\mathbf{E}^T=\mathbf{G}(\bar{\varsigma})+\mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T\preceq 0.$$

This means that $\bar{\boldsymbol{x}}$ is also a local maximizer of Problem $(\mathcal{P})$ under the assumption (26), i.e., there exists a neighborhood $\mathcal{X}_o\times\mathcal{S}_o\subset\mathbb{R}^n\times\mathcal{S}_a^-$ of $(\bar{\boldsymbol{x}},\ \bar{\boldsymbol{\varsigma}})$ such that

$$\max_{\boldsymbol{x}\in\mathcal{X}_o}\Pi(\boldsymbol{x})=\Xi(\bar{\boldsymbol{x}},\bar{\varsigma})=\max_{\boldsymbol{\varsigma}\in\mathcal{S}_o}\Pi^d(\boldsymbol{\varsigma}).$$

Finally, we can show, in a similar way, that if $\bar{\boldsymbol{x}}=[\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$ is a local maximizer of Problem $(\mathcal{P})$ and $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$, the corresponding $\bar{\boldsymbol{\varsigma}}$ is also a local maximizer of Problem $(\mathcal{P}^d)$. Therefore, the tri-duality theorem is proved. ■

Remark 4. The strong triality Theorem 3.1 can also be used to identify saddle points of the primal problem, i.e. $\bar{\varsigma}\in\mathcal{S}_a^-$ is a saddle point of $\Pi^d(\boldsymbol{\varsigma})$ if and only if $\bar{\boldsymbol{x}}=\mathbf{G}(\bar{\varsigma})^{-1}\boldsymbol{f}$ is a saddle point of $\Pi(\boldsymbol{x})$ on $\mathcal{X}_a$. Since the saddle points do not produce computational difficulties in numerical optimization, and do not exist physically in static systems, these points are excluded from the triality theory.

Remark 5. By the proof of Theorem 3.1, we know that if there exists a critical point $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$ such that $\bar{\boldsymbol{\varsigma}}$ is a local minimizer of Problem $(\mathcal{P}^d)$, then $\mathbf{F}(\bar{\boldsymbol{x}})$ must be invertible. On the other hand, if the symmetric matrices $\{\boldsymbol{B}^k\}$ are linearly dependent, then $\mathbf{F}(\boldsymbol{x})$ is not invertible for any $\boldsymbol{x}\in\mathbb{R}^n$. In this case, the corresponding canonical dual problem $(\mathcal{P}^d)$ has no local minimizers in $\mathcal{S}_a^-$, and for any critical point $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$, the analytical solution $\bar{\boldsymbol{x}}=[\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$ is not a local minimizer of $\Pi(\boldsymbol{x})$.

4. Refined Triality Theory for General Quartic Polynomial Optimization.

Let us recall the primal problem and its canonical dual problem in the general quartic polynomial case ($n\neq m$):

$$(\mathcal{P}):\quad \text{ext}\quad\left\{\Pi(\boldsymbol{x})=\frac{1}{2}\sum_{k=1}^{m}\left(\frac{1}{2}\boldsymbol{x}^T\mathbf{B}^k\boldsymbol{x}-d^k\right)^2+\frac{1}{2}\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}-\boldsymbol{x}^T\boldsymbol{f}\mid\boldsymbol{x}\in\mathbb{R}^n\right\}, \tag{44}$$

$$\left(\mathcal{P}^d\right):\quad \text{ext}\quad\left\{\Pi^d(\boldsymbol{\varsigma})=-\frac{1}{2}\boldsymbol{f}^T[\mathbf{G}(\boldsymbol{\varsigma})]^{-1}\boldsymbol{f}-\frac{1}{2}\boldsymbol{\varsigma}^T\boldsymbol{\varsigma}-\boldsymbol{\varsigma}^T\boldsymbol{d}\mid\boldsymbol{\varsigma}\in\mathcal{S}_a\subset\mathbb{R}^m\right\}. \tag{45}$$

Suppose that $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{\varsigma}}$ are the critical points of Problem $(\mathcal{P})$ and Problem $(\mathcal{P}^d)$, respectively, where $\bar{\boldsymbol{x}}=[\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$. It is easy to verify that

$$\nabla^2\Pi(\bar{\boldsymbol{x}})=\mathbf{G}(\bar{\varsigma})+\mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T\in\mathbb{R}^{n\times n} \tag{46}$$

$$\nabla^2\Pi^d(\bar{\varsigma})=-\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}})-\mathbf{I}\in\mathbb{R}^{m\times m}. \tag{47}$$

In this case,

$$\mathbf{F}(\boldsymbol{x})=\left[\mathbf{B}^1\boldsymbol{x},\mathbf{B}^2\boldsymbol{x},\cdots,\mathbf{B}^m\boldsymbol{x}\right]\in\mathbb{R}^{n\times m}.$$

To continue, we show the following lemmas.

Lemma 4.1. *Suppose that $m<n$. Let the critical point $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$ be a local minimizer of $\Pi^d(\boldsymbol{\varsigma})$, and let $\bar{\boldsymbol{x}}=[\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$. Then, there exists a matrix $\mathbf{P}\in\mathbb{R}^{n\times m}$ with* $\text{rank}(\mathbf{P})=\text{m}$ *such that*

$$\mathbf{P}^T\nabla^2\Pi(\bar{\boldsymbol{x}})\mathbf{P}\succeq\mathbf{0}. \tag{48}$$

Proof. By the fact that the critical point $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$ is a local minimizer of $\Pi^d(\boldsymbol{\varsigma})$, we have $\nabla\Pi^d(\bar{\varsigma})=0$ and $\nabla^2\Pi^d(\bar{\varsigma})\succeq 0$. It follows that

$$-\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}})\succeq\mathbf{I}\in\mathbb{R}^{m\times m}.$$

Thus, $\text{rank}(\mathbf{F}(\bar{\boldsymbol{x}}))=\text{m}$. Since $\bar{\boldsymbol{\varsigma}}\in\mathcal{S}_a^-$ and $\mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T\succeq 0$, there exists a non-singular matrix $\mathbf{T}\in\mathbb{R}^{n\times n}$ such that

$$\mathbf{T}^T\mathbf{G}(\bar{\varsigma})\mathbf{T}=\text{Diag}\ (-\lambda_1,\cdots,-\lambda_n) \tag{49}$$

and

$$\mathbf{T}^T\mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T\mathbf{T}=\text{Diag}\ (a_1,\cdots,a_m,0,\ldots,0), \tag{50}$$

where $\lambda_i>0,\ i=1,\cdots,n$, and $a_j>0,\ j=1,\cdots,m$.

According to the singular value decomposition theory [22], there exist orthogonal matrices $\mathbf{U}$ and $\mathbf{E}$ such that

$$\mathbf{T}^T\mathbf{F}(\bar{\boldsymbol{x}}) = \mathbf{U}\begin{pmatrix} \sqrt{a_1} & & \\ & \ddots & \\ & & \sqrt{a_m} \\ 0 & \cdots & 0 \\ & \cdots & \\ 0 & \cdots & 0 \end{pmatrix}\mathbf{E}.$$

Therefore, $\mathbf{U}$ is an identity matrix. Let

$$\mathbf{R} = \begin{pmatrix} \sqrt{a_1} & & \\ & \ddots & \\ & & \sqrt{a_m} \\ 0 & \cdots & 0 \\ & \cdots & \\ 0 & \cdots & 0 \end{pmatrix}.$$

Then,

$$\begin{aligned} \nabla^2\Pi^d(\bar{\varsigma}) &= -\mathbf{F}(\bar{\boldsymbol{x}})^T[\mathbf{G}(\bar{\varsigma})]^{-1}\mathbf{F}(\bar{\boldsymbol{x}}) - \mathbf{I} \\ &= -\left(\mathbf{F}^T\mathbf{T}\right)\left[\mathbf{T}^T\mathbf{G}(\bar{\varsigma})\mathbf{T}\right]^{-1}\mathbf{T}^T\mathbf{F}(\bar{\boldsymbol{x}}) - \mathbf{I} \\ &= -\mathbf{E}^T\mathbf{R}\mathbf{U}^T[\text{Diag }(-\lambda_1,\cdots,-\lambda_n)]^{-1}\mathbf{U}\mathbf{R}\mathbf{E} - \mathbf{I} \in \mathbb{R}^{m\times m}. \end{aligned}$$

Since $\nabla^2\Pi^d(\bar{\varsigma}) \succeq 0$, $\mathbf{U}$ is an identity matrix, and $\mathbf{E}$ is an orthogonal matrix, we have

$$-\mathbf{R}[\text{Diag }(\lambda_1,\cdots,\lambda_n)]^{-1}\mathbf{R} - \mathbf{I}_{m\times m} = \text{ Diag }\left(\frac{a_1}{\lambda_1} - 1,\cdots,\frac{a_m}{\lambda_m} - 1\right) \succeq 0.$$

Thus, $a_i \geq \lambda_i$, $i = 1,\cdots,m$. Note that

$$\mathbf{T}^T\nabla^2\Pi(\bar{\boldsymbol{x}})\mathbf{T} = \text{Diag }(a_1-\lambda_1,\cdots,a_m-\lambda_m,-\lambda_{m+1},\cdots,-\lambda_n). \tag{51}$$

Let $\mathbf{J} = [\mathbf{I}_{m\times m}, \mathbf{0}_{m\times(n-m)}]^T$. Then, we have

$$\mathbf{J}^T\mathbf{T}^T\nabla^2\Pi(\bar{\boldsymbol{x}})\mathbf{T}\mathbf{J} = \text{Diag }(a_1-\lambda_1,\cdots,a_m-\lambda_m) \succeq \mathbf{0}. \tag{52}$$

Let $\mathbf{P} = \mathbf{T}\mathbf{J}$. Clearly, $\text{rank}(\mathbf{P}) = \text{m}$ and $\mathbf{P}^T\nabla^2\Pi(\bar{\boldsymbol{x}})\mathbf{P} = \text{Diag }(a_1-\lambda_1,\cdots,a_m-\lambda_m) \succeq \mathbf{0}$. The proof is completed. ∎

In a similar way, we can prove the following lemma.

Lemma 4.2. *Suppose that $m > n$. Let $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$ be a critical point, which is a local minimizer of Problem $(\mathcal{P})$, where $\bar{\varsigma} \in \mathcal{S}_a^-$. Then, there exists a matrix $\mathbf{Q} \in \mathbb{R}^{m\times n}$ with $\text{rank}(\mathbf{Q}) = \text{n}$ such that*

$$\mathbf{Q}^T\nabla^2\Pi^d(\bar{\varsigma})\mathbf{Q} \succeq \mathbf{0}. \tag{53}$$

Let $\mathbf{p}_1,\cdots,\mathbf{p}_m$ be the m column vectors of $\mathbf{P}$ and let $\mathbf{q}_1,\cdots,\mathbf{q}_n$ be the n column vectors of $\mathbf{Q}$, respectively. Clearly, $\mathbf{p}_1,\cdots,\mathbf{p}_m$ are m independent vectors and $\mathbf{q}_1,\cdots,\mathbf{q}_n$ are n independent vectors. We introduce the following two subspaces

$$\mathcal{X}_\flat = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \bar{\boldsymbol{x}} + \theta_1\mathbf{p}_1 + \cdots + \theta_m\mathbf{p}_m,\ \theta_i \in \mathbb{R},\ i = 1,\cdots,m\}, \tag{54}$$

$$\mathcal{S}_\flat = \{\varsigma \in \mathbb{R}^m \mid \varsigma = \bar{\varsigma} + \vartheta_1\mathbf{q}_1 + \cdots + \vartheta_n\mathbf{q}_n,\ \vartheta_i \in \mathbb{R},\ i = 1,\cdots,n\}. \tag{55}$$

Theorem 4.3 (Refined Triality Theorem)**.**

Suppose that the assumption (26) is satisfied. Let $\bar{\varsigma}$ be a critical point of $\Pi^d(\varsigma)$ and let $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\varsigma})]^{-1}\boldsymbol{f}$.

If $\bar{\varsigma} \in \mathcal{S}_a^+$, then the canonical min-max duality holds in the strong form:

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x}\in\mathbb{R}^n}\Pi(\boldsymbol{x}) \Longleftrightarrow \max_{\varsigma\in\mathcal{S}_a^+}\Pi^d(\varsigma) = \Pi^d(\bar{\varsigma}). \tag{56}$$

If $\bar{\varsigma} \in \mathcal{S}_a^-$, then there exists a neighborhood $\mathcal{X}_o \times \mathcal{S}_o \subset \mathbb{R}^n \times \mathcal{S}_a^-$ of $(\bar{\boldsymbol{x}},\bar{\varsigma})$ such that the double-max duality holds in the strong form

$$\Pi(\bar{\boldsymbol{x}}) = \max_{\boldsymbol{x}\in\mathcal{X}_o}\Pi(\boldsymbol{x}) \Longleftrightarrow \max_{\varsigma\in\mathcal{S}_o}\Pi^d(\varsigma) = \Pi^d(\bar{\varsigma}). \tag{57}$$

However, the double-min duality statement holds conditionally in the following super-symmetrical forms.

1. *If $m < n$ and $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ is a local minimizer of $\Pi^d(\boldsymbol{\varsigma})$, then $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f}$ is a saddle point of $\Pi(\boldsymbol{x})$ and the double-min duality holds weakly on $\mathcal{X}_o \cap \mathcal{X}_\flat \times \mathcal{S}_o$, i.e.*

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x} \in \mathcal{X}_o \cap \mathcal{X}_\flat} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{\varsigma} \in \mathcal{S}_o} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}); \tag{58}$$

2. *If $m > n$ and $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f}$ is a local minimizer of $\Pi(\boldsymbol{x})$, then $\bar{\boldsymbol{\varsigma}}$ is a saddle point of $\Pi^d(\boldsymbol{\varsigma})$ and the double-min duality holds weakly on $\mathcal{X}_o \times \mathcal{S}_o \cap \mathcal{S}_\flat$, i.e.*

$$\Pi(\bar{\boldsymbol{x}}) = \min_{\boldsymbol{x} \in \mathcal{X}_o} \Pi(\boldsymbol{x}) = \min_{\boldsymbol{\varsigma} \in \mathcal{S}_o \cap \mathcal{S}_\flat} \Pi^d(\boldsymbol{\varsigma}) = \Pi^d(\bar{\boldsymbol{\varsigma}}). \tag{59}$$

Proof. The proof of the statements (56) and (57) are similar to that given for the proof of Theorem 3.1. Thus, it suffices to prove the double-min duality statements (58) and (59).

Firstly, we suppose that $m < n$ and $\bar{\boldsymbol{\varsigma}}$ is a local minimizer of Problem $(\mathcal{P}^d)$. Define

$$\varphi(t_1, \cdots, t_m) = \Pi(\bar{\boldsymbol{x}} + t_1\bar{\boldsymbol{x}}_1 + \cdots + t_m\bar{\boldsymbol{x}}_m). \tag{60}$$

From (47), we obtain

$$-\mathbf{F}(\bar{\boldsymbol{x}})^T [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1} \mathbf{F}(\bar{\boldsymbol{x}}) \succeq \mathbf{I} \in \mathbb{R}^{m\times m}.$$

Thus, $\mathbf{F}(\bar{\boldsymbol{x}})^T [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1} \mathbf{F}(\bar{\boldsymbol{x}})$ is a non-singular matrix and rank $(\mathbf{F}(\bar{\boldsymbol{x}})) = m < n$. We claim that $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f}$ is not a local minimizer of Problem $(\mathcal{P})$. On a contrary, suppose that $\bar{\boldsymbol{x}}$ is also a local minimizer. Then, we have

$$\nabla^2\Pi(\bar{\boldsymbol{x}}) = \mathbf{G}(\bar{\boldsymbol{\varsigma}}) + \mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T \succeq 0.$$

Thus, $\mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T \succeq -\mathbf{G}(\bar{\boldsymbol{\varsigma}})$. Since $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$ and rank $(\mathbf{F}) = m$, it is clear that

$$n = \text{rank } (\mathbf{G}(\bar{\boldsymbol{\varsigma}})) = \text{rank } \left(\mathbf{F}(\bar{\boldsymbol{x}})\mathbf{F}(\bar{\boldsymbol{x}})^T\right) = m.$$

This is a contradiction. Therefore, $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f}$ is a saddle point of Problem $(\mathcal{P})$.

It is easy to verify that $\Pi(\bar{\boldsymbol{x}}) = \Pi^d(\bar{\boldsymbol{\varsigma}})$. Thus, to prove (58), it suffices to prove that $\mathbf{0} \in \mathbb{R}^m$ is a local minimizer of the function $\varphi(t_1, \cdots, t_n)$.

It is easy to verify that

$$\nabla\varphi(0, \cdots, 0) = [(\nabla\Pi(\bar{\boldsymbol{x}}))^T\mathbf{p}_1, \cdots, (\nabla\Pi(\bar{\boldsymbol{x}}))^T\mathbf{p}_m]^T = \nabla\Pi(\bar{\boldsymbol{x}}))^T\mathbf{P} = \mathbf{0} \tag{61}$$

and

$$\nabla^2\varphi(0, \cdots, 0) = \mathbf{P}^T\nabla^2\Pi(\bar{\boldsymbol{x}})\mathbf{P}. \tag{62}$$

In light of Lemma 4.1 and the assumption (26), it follows that $\mathbf{0} \in \mathbb{R}^m$ is, indeed, a local minimizer of the function $\varphi(t_1, \cdots, t_m)$.

In a similar way, we can establish the case of $m > n$. The proof is completed. ∎

Remark 6. Theorem 4.3 shows that both the canonical min-max and double-max duality statements hold strongly for general cases; the double-min duality holds strongly for $n = m$ but weakly for $n \neq m$ in a symmetrical form. The "certain additional conditions" are simply the intersection $\mathcal{X}_o \bigcap \mathcal{X}_\flat$ for $m < n$ and $\mathcal{S}_o \bigcap \mathcal{S}_\flat$ for $m > n$. Therefore, the open problem left in 2003 [9, 10] is solved for the double-well potential function $W(\boldsymbol{x})$.

The triality theory has been challenged recently in a series of more than seven papers, see, for example, [31, 32]. In the first version of [32], Voisei and Zalinescu wrote: "we consider that it is important to point out that the main results of this (triality) theory are false. This is done by providing elementary counter-examples that lead to think that a correction of this theory is impossible without falling into trivia". It turns out that most of these counter-examples simply use the double-well function $W(\boldsymbol{x})$ with $n \neq m$. In fact, these counter-examples address the same type of open problem for the double-min duality left unaddressed in [9, 10]. Indeed, by Theorem 4.3, we know that both the canonical min-max duality and the double-max duality hold strongly for the general case $n \neq m$. However, based on the weak double-min duality, one can easily construct other *V-Z type counterexamples*, where the strong double-min duality holds conditionally when $n \neq m$. Also, interested readers should find that the references [9, 10] never been cited in any one of their papers.

5. **Numerical Experiments.** In this section, some simple numerical examples are presented to illustrate the canonical duality theory.

Example 1 ($m = n = 2$). Let us first consider Problem ($\mathcal{P}$) with $n = m = 2$.

$$\text{ext}\left\{\Pi(\boldsymbol{x}) = \frac{1}{2}\left[\left(\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{B}^{1}\boldsymbol{x} - \mathrm{d}_1\right)^2 + \left(\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{B}^{2}\boldsymbol{x} - \mathrm{d}_2\right)^2\right] + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{A}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{f} \,|\, \boldsymbol{x} \in \mathbb{R}^2\right\}, \quad (63)$$

where

$$\mathbf{A} = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}, \ \mathbf{B}^1 = \begin{bmatrix} b_1 & 0 \\ 0 & 0 \end{bmatrix}, \ \mathbf{B}^2 = \begin{bmatrix} 0 & 0 \\ 0 & b_2 \end{bmatrix}, \ \boldsymbol{f} = [f_1, f_2]^T.$$

The canonical dual problem can be expressed as

$$\Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\left(\frac{f_1^2}{a_1 + \varsigma_1 b_1} + \frac{f_2^2}{a_2 + \varsigma_2 b_2}\right) - \frac{1}{2}\left(\varsigma_1^2 + \varsigma_2^2\right) - (d_1\varsigma_1 + d_2\varsigma_2).$$

Thus,

$$\nabla\Pi^d(\boldsymbol{\varsigma}) = \begin{bmatrix} \frac{b_1 f_1^2}{2(a_1+\varsigma_1 b_1)^2} - \varsigma_1 - 1 \\ \frac{b_2 f_2^2}{2(a_2+\varsigma_2 b_2)^2} - \varsigma_2 - 1 \end{bmatrix}$$

and

$$\nabla^2\Pi^d(\boldsymbol{\varsigma}) = \begin{bmatrix} -b_1^2 f_1^2 (a_1 + \varsigma_1 b_1)^{-3} - 1 & 0 \\ 0 & -b_2^2 f_2^2 (a_2 + \varsigma_2 b_2)^{-3} - 1 \end{bmatrix}.$$

Now, we take $b_1 = b_2 = f_1 = f_2 = 1$, $d_1 = d_2 = 1$ and $a_1 = -2$, $a_2 = -3$. It is easy to check that $\Pi^d(\boldsymbol{\varsigma})$ has only one critical point $\bar{\boldsymbol{\varsigma}}_1 = \left(1 + \sqrt{2}, 3.35991198\right)$ in $\mathcal{S}_a^+$ and four critical points $\bar{\boldsymbol{\varsigma}}_2 = (3/2, 2.59827880)$, $\bar{\boldsymbol{\varsigma}}_3 = \left(1 - \sqrt{2}, -0.45819078\right)$, $\bar{\boldsymbol{\varsigma}}_4 = \left(1 - \sqrt{2}, 2.59827880\right)$, $\bar{\boldsymbol{\varsigma}}_5 = (3/2, -0.45819078)$ in S_a^-, respectively. Furthermore, $\bar{\boldsymbol{\varsigma}}_2$ is a local minimizer and $\bar{\boldsymbol{\varsigma}}_3$ is a local maximizer; the solutions $\bar{\boldsymbol{\varsigma}}_4$ and $\bar{\boldsymbol{\varsigma}}_5$ are saddle points of $\Pi^d(\boldsymbol{\varsigma})$ in $\mathcal{S}_a^-$. Thus, by Theorem 3.1, we know that $\bar{\boldsymbol{x}}_1 = (2.41421356, 2.77845711)$ is a global minimizer, while $\bar{\boldsymbol{x}}_2 = (-2, -2.48928859)$ is a local minimizer and $\bar{\boldsymbol{x}}_3 = (-0.41421356, -0.28916855)$ is a local maximizer. The corresponding values of the cost function are

$$\Pi(\bar{\boldsymbol{x}}_1) = -14.0421 < \Pi(\bar{\boldsymbol{x}}_2) = -4.3050 < \Pi(\bar{\boldsymbol{x}}_3) = 0.5971.$$

The graph of $\Pi(x)$ and its contour are depicted in Figure 1.

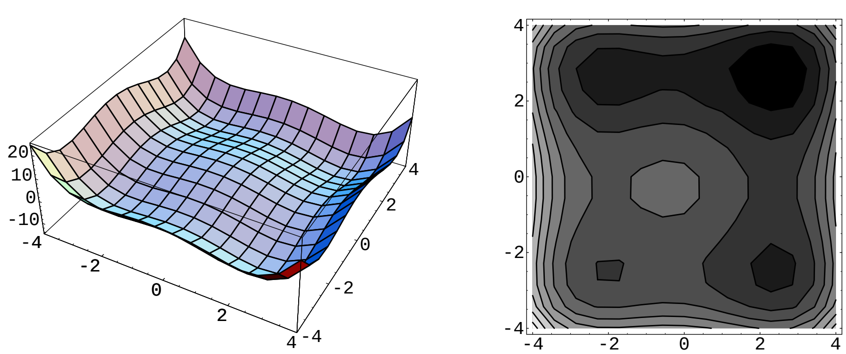

FIGURE 1. Graph of $\Pi(\boldsymbol{x})$ (left) and contours of $\Pi(\boldsymbol{x})$ (right) for Example 1

Example 2 ($n = 2, m = 1$). We now consider Problem ($\mathcal{P}$) with $n = 2$, $m = 1$, $A = \text{Diag}(-0.2, -0.8)$, $B = \text{Diag}(1, 1)$, $\boldsymbol{f} = (0.9, 0.3)^T$, $d = 4$. Then, its dual problem is

$$\Pi^d(\varsigma) = -\frac{1}{2}\left(\frac{0.9^2}{-0.2 + \varsigma} + \frac{0.3^2}{-0.8 + \varsigma}\right) - \frac{1}{2}\varsigma^2 - 4\varsigma.$$

We can verify that $\Pi^d(\varsigma)$ has one critical point $\bar{\varsigma}_1 = -0.90489505$ in S_a^+ and two critical points $\bar{\varsigma}_2 = -0.12552589$ and $\bar{\varsigma}_3 = -3.974788888$ in S_a^-. Furthermore, $\bar{\varsigma}_2$ is a local minimizer and $\bar{\varsigma}_3$ is a local maximizer of $\Pi^d(\varsigma)$. According to Theorem 4.3, $\bar{\boldsymbol{x}}_1 = (A + \bar{\varsigma}_1 B)^{-1}\boldsymbol{f} = (1.27678581, 2.86000142)^T$ is the unique global minimizer of $\Pi(\boldsymbol{x})$, $\bar{\boldsymbol{x}}_2 = (A + \bar{\varsigma}_2 B)^{-1}\boldsymbol{f} = (-2.76475703, -0.32414004)^T$ is

a saddle point and $\bar{\boldsymbol{x}}_3 = (A + \bar{\varsigma}_3 B)^{-1}\boldsymbol{f} = (-0.21557976, -0.06283003)^T$ is a local maximizer of $\Pi(\boldsymbol{x})$. Let $\boldsymbol{p} = (1,0)^T$ and $\varphi(\theta) = \Pi(\bar{\boldsymbol{x}}_2 + \theta\boldsymbol{p})$. Then, it is easy to verify that there exists neighborhoods $\mathcal{X}_0 \subset \mathbb{R}$ and $\mathcal{S}_0 \subset \mathbb{R}$ such that $0 \in \mathcal{X}_0$, $\bar{\varsigma}_2 \in \mathcal{S}_0$ and

$$\min_{\theta\in\mathcal{X}_0} \varphi(\theta) = \min_{\varsigma\in\mathcal{S}_0} \Pi^d(\varsigma).$$

This example shows that even if $n > m$, the canonical min-max duality and the double-max duality still hold strongly. However, the double-min duality statement should be refined into an $m-$ dimensional subspace in this case.

Example 3 ($n = 1, m = 2$). We now consider Problem $(\mathcal{P})$ with $n = 1$, $m = 2$, $A = -0.2$, $B^1 = 0.3$, $B^2 = 0.7$, $d_1 = 3$, $d_2 = 2.7$ and $f = 1.4$. Then, its dual problem is

$$\Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\left(\frac{f^2}{A + \varsigma_1 B^1 + \varsigma_2 B^2}\right) - \frac{1}{2}(\varsigma_1^2 + \varsigma_2^2) - (d_1\varsigma_1 + d_2\varsigma_2).$$

We can verify that $\Pi^d(\boldsymbol{\varsigma})$ has one critical point $\bar{\boldsymbol{\varsigma}}_1 = (-0.35012607, 3.48303916)^T$ in S_a^+ and two critical points $\bar{\boldsymbol{\varsigma}}_2 = (-2.98705125, -2.66978626)^T$ and $\bar{\boldsymbol{\varsigma}}_3 = (-0.70765026, 2.64881606)^T$ in S_a^-. Furthermore, $\bar{\boldsymbol{\varsigma}}_2$ is a local maximizer and $\bar{\boldsymbol{\varsigma}}_3$ is a saddle point of $\Pi^d(\boldsymbol{\varsigma})$. According to Theorem 4.3, $\bar{\boldsymbol{x}}_1 = G(\bar{\boldsymbol{\varsigma}}_1)^{-1} f = 4.20307342$ is the unique global minimizer of $\Pi(\boldsymbol{x})$, $\bar{\boldsymbol{x}}_2 = G(\bar{\boldsymbol{\varsigma}}_2)^{-1} f = -0.29381114$ is a local maximizer of $\Pi(\boldsymbol{x})$. Let $\boldsymbol{q} = (1,0)^T$ and $\psi(\vartheta) = \Pi^d(\bar{\boldsymbol{\varsigma}}_3 + \vartheta\boldsymbol{q})$. Then, it is easy to verify that there exists neighborhoods $\mathcal{X}_0 \subset \mathbb{R}$ and $\mathcal{S}_0 \subset \mathbb{R}$ such that $\bar{\boldsymbol{x}}_3 \in \mathcal{X}_0$, $0 \in \mathcal{S}_0$ and

$$\min_{\boldsymbol{x}\in\mathcal{X}_0} \Pi(\boldsymbol{x}) = \min_{\vartheta\in\mathcal{S}_0} \psi(\vartheta).$$

This example shows that if $n < m$, the canonical min-max duality and the double-max duality still hold strongly. However, the double-min duality statement should be refined into an $n-$ dimensional subspace in this case.

Example 4 Linear Perturbation. Let us consider the following optimization problem without input ($\boldsymbol{f} = 0$)

$$(\mathcal{P}_2) : \text{ext}\left\{\Pi(\boldsymbol{x}) = \frac{1}{2}\left[\left(\frac{1}{2}(\mathrm{x}_1 + \mathrm{x}_2)^2 - \frac{1}{2}\right)^2 + \left(\frac{1}{2}(\mathrm{x}_1 - \mathrm{x}_2) - \frac{1}{2}\right)^2\right] \mid \boldsymbol{x} \in \mathbb{R}^2\right\}.$$

Problem $(\mathcal{P}_2)$ has four global minimizers $\bar{\boldsymbol{x}}_1 = (1,0)$, $\bar{\boldsymbol{x}}_2 = (0,-1)$, $\bar{\boldsymbol{x}}_3 = (0,1)$, $\bar{\boldsymbol{x}}_4 = (-1,0)$ and the optimal cost value is 0. Its canonical dual problem is

$$\Pi^d(\boldsymbol{\varsigma}) = -\frac{1}{2}\left(\varsigma_1^2 + \varsigma_2^2\right) - (\varsigma_1 + \varsigma_2).$$

$\Pi^d(\boldsymbol{\varsigma})$ has only one critical point $\bar{\boldsymbol{\varsigma}} = \left(-\frac{1}{2}, -\frac{1}{2}\right) \in \mathcal{S}_a^-$. Furthermore, we can check that $\bar{\boldsymbol{x}} = [G(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f} = (0,0)$ is a local maximizer of Problem $(\mathcal{P}_2)$. Thus, we cannot use the canonical dual transformation method to obtain the global minimizer of Problem $(\mathcal{P}_2)$ since this problem is in a perfect symmetrical form without input, which allows more than one global minimizer. Now we perturb Problem $(\mathcal{P}_2)$ as follows.

$$\left(\mathcal{P}_2^b\right) : \text{ext}_{\boldsymbol{x}\in\mathbb{R}^2} \Pi_2(\boldsymbol{x}) = \frac{1}{2}\left[\left(\frac{1}{2}(\mathrm{x}_1 + \mathrm{x}_2)^2 - \frac{1}{2}\right)^2 + \left(\frac{1}{2}(\mathrm{x}_1 - \mathrm{x}_2)^2 - \frac{1}{2}\right)^2\right] - (\mathrm{x}_1\mathrm{f}_1 + \mathrm{x}_2\mathrm{f}_2).$$

Its canonical dual function is expressed as

$$\Pi_2^d(\boldsymbol{\varsigma}) = -\frac{1}{8\varsigma_1\varsigma_2}\left[(\varsigma_1 + \varsigma_2)\left(f_1^2 + f_2^2\right) + 2(\varsigma_1 - \varsigma_2) f_1 f_2\right] - \frac{1}{2}\left(\varsigma_1^2 + \varsigma_2^2\right) - \frac{1}{2}(\varsigma_1 + \varsigma_2).$$

Taking $f_1 = 0.001$, $f_2 = 0.005$ and solving $\nabla\Pi_2^d(\bar{\boldsymbol{\varsigma}}) = 0$, the results obtained are listed in Table 1. We can see that $\bar{\boldsymbol{\varsigma}} = (0.00299107, 0.00199602) \in \mathcal{S}_a^+$ and $\Pi_2^d(\boldsymbol{\varsigma}) = -0.00500648$. Thus, $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f} = (0.000495793, 1.00249)$ is the global minimizer of Problem $\left(\mathcal{P}_2^b\right)$. Clearly, this $\bar{\boldsymbol{x}}$ is very close to $\bar{\boldsymbol{x}}_3$. If we take $f_1 = 0.001$, $f_2 = -0.005$, the global minimizer of Problem $\left(\mathcal{P}_2^b\right)$ is $\bar{\boldsymbol{x}} = [\mathbf{G}(\bar{\boldsymbol{\varsigma}})]^{-1}\boldsymbol{f} = (0.000496288, -1.00249)$ which is close to $\bar{\boldsymbol{x}}_2 = (0,-1)$. This example shows that if the canonical dual problem has no critical point in $\mathcal{S}_a^+$, a linear perturbation could be used to solve the primal problem.

TABLE 1. Numerical results for Example 3

$\boldsymbol{\varsigma} = (\varsigma_1, \varsigma_2)$	$\boldsymbol{x} = (x_1, x_2)$	$G(\boldsymbol{\varsigma})$	The primal problem
$(-0.49982, -0.499992)$	$(-0.00100008, -0.00500014)$	$G \prec \mathbf{0}$	local max
$(-0.00300907, -0.499992)$	$(-0.496493, -0.500493)$	$G \prec \mathbf{0}$	saddle point
$(0.00299107, -0.499992)$	$(-0.499493, -0.503493)$	$G \prec \mathbf{0}$	saddle point
$(-0.49982, -0.00200402)$	$(0.495997, -0.501997)$	$G \prec \mathbf{0}$	saddle point
$(-0.00300907, -0.00200402)$	$(0.000504216, -0.99749)$	$G \prec \mathbf{0}$	local min
$(0.00299107, -0.00200402)$	$(1.00049, 0.00249576)$	indefinite	saddle point
$(-0.49982, 0.00199602)$	$(-0.503997, 0.497997)$	indefinite	local max
$(-0.00300907, 0.00199602)$	$(-0.99949, 0.00250409)$	indefinite	saddle point
$(0.00299107, 0.00199602)$	$(0.000495793, 1.00249)$	$G \succ \mathbf{0}$	global min

6. **Appendix.** In this Appendix, we present several lemmas which are needed for the proofs of Theorem 3.1 and Theorem 4.3.

Lemma 6.1 (Singular value decomposition [22])**.** *For any given* $\mathbf{G} \in \mathbb{R}^{n\times n}$ *with* $\text{rank}(\mathbf{G}) = \text{r}$*, there exist* $\mathbf{U} \in \mathbb{R}^{n\times n}$*,* $\mathbf{D} \in \mathbb{R}^{n\times n}$ *and* $\mathbf{R} \in \mathbb{R}^{n\times n}$ *such that*

$$\mathbf{G} = \mathbf{UDR},$$

where $\mathbf{U}$*,* $\mathbf{R}$ *are orthogonal matrices, i.e.,* $\mathbf{U}^T\mathbf{U} = \mathbf{I} = \mathbf{R}^T\mathbf{R}$*, and* $\mathbf{D} = Diag\,(\sigma_1, \cdots, \sigma_r, 0, \cdots, 0)$*,* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ *.*

Lemma 6.2. *Suppose that* $\mathbf{G}$ *and* $\mathbf{U}$ *are positive definite. Then,* $\mathbf{G} \succeq \mathbf{U}$ *if and only if* $\mathbf{U}^{-1} \succeq \mathbf{G}^{-1}$*.*

Proof. The proof is trivial and is omitted here.

Lemma 6.3 (Proposition 2.1 in [21])**.** *For any given symmetric matrix* $\mathbf{M}$ *expressed in the form*

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix}$$

such that $\mathbf{M}_{22} \succ 0$*. Then,* $\mathbf{M} \succeq 0$ *if and only if* $\mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{-1}\mathbf{M}_{21} \succeq 0$*.*

The following lemma plays a key role in the proof of Theorem 3.1 and Theorem 4.3.

Lemma 6.4. *Suppose that* $\mathbf{P} \in \mathbb{R}^{n\times n}$*,* $\mathbf{U} \in \mathbb{R}^{m\times m}$*, and* $\mathbf{D} \in \mathbb{R}^{n\times m}$*. Furthermore,*

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{0}_{(n-r)\times(m-r)} \end{bmatrix} \in \mathbb{R}^{n\times n},$$

where $\mathbf{D}_{11} \in \mathbb{R}^{r\times r}$ *is nonsingular,* $r = \text{rank}(\mathbf{D})$*, and*

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} \prec 0, \quad \mathbf{U} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{0}_{r\times(m-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{U}_{22} \end{bmatrix} \succ 0,$$

$\mathbf{P}_{ij}$ *and* $\mathbf{U}_{ii}$*,* $i, j = 1, 2$*, are of appropriate dimension matrices. Then,*

$$\mathbf{P} + \mathbf{DUD}^T \preceq 0 \Longleftrightarrow -\mathbf{D}^T\mathbf{P}^{-1}\mathbf{D} - \mathbf{U}^{-1} \preceq 0. \tag{64}$$

Proof. Suppose that $\mathbf{P} + \mathbf{DUD}^T \preceq 0$. Then,

$$-\mathbf{P} - \mathbf{DUD}^T = \begin{bmatrix} -\mathbf{P}_{11} - \mathbf{D}_{11}\mathbf{U}_{11}\mathbf{D}_{11}^T & -\mathbf{P}_{12} \\ -\mathbf{P}_{21} & -\mathbf{P}_{22} \end{bmatrix} \succeq 0.$$

Since $\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} \prec 0$, it follows that $-\mathbf{P}_{22} \succ 0$. By Lemma 6.3, we have the following inequality

$$-\mathbf{P}_{11} - \mathbf{D}_{11}\mathbf{U}_{11}\mathbf{D}_{11}^T + \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21} \succeq 0 \tag{65}$$

which leads to
$$-\mathbf{P}_{11}+\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\succeq \mathbf{D}_{11}\mathbf{U}_{11}\mathbf{D}_{11}^{T}.$$
Since $\mathbf{P}\prec 0$ and $\mathbf{U}\succ 0$, it follows from Lemma 6.2 that
$$\left(-\mathbf{P}_{11}+\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\right)^{-1}\preceq \left(\mathbf{D}_{11}^{T}\right)^{-1}\mathbf{U}_{11}^{-1}\mathbf{D}_{11}^{-1}.$$
Thus,
$$\mathbf{D}_{11}^{T}\left(-\mathbf{P}_{11}+\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\right)^{-1}\mathbf{D}_{11}\preceq \mathbf{U}_{11}^{-1}. \tag{66}$$
Note that
$$\mathbf{P}^{-1}=\begin{bmatrix} \left(\mathbf{P}_{11}-\mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\right)^{-1} & \mathbf{P}_{11}^{-1}\mathbf{P}_{12}\left(\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}-\mathbf{P}_{22}\right)^{-1} \\ \left(\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}-\mathbf{P}_{22}\right)^{-1}\mathbf{P}_{21}\mathbf{P}_{11}^{-1} & \left(\mathbf{P}_{22}-\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\right) \end{bmatrix}.$$
By virtue of Lemma 6.3, we obtain
$$-\begin{bmatrix} \mathbf{D}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}^{T}\mathbf{P}^{-1}\begin{bmatrix} \mathbf{D}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\preceq \begin{bmatrix} \mathbf{U}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{22}^{-1} \end{bmatrix}=\mathbf{U}^{-1},$$
i.e., the right hand side of (64) holds.

In a similar way, we can show that if $-\mathbf{D}^{T}\mathbf{P}^{-1}\mathbf{D}-\mathbf{U}^{-1}\preceq 0$, then $\mathbf{P}+\mathbf{D}\mathbf{U}\mathbf{D}^{T}\preceq 0$. The proof is thus completed. ∎

7. **Conclusion Remarks.** In this paper, we presented a rigorous proof of the double-min duality in the triality theory for a quartic polynomial optimization problem based on elementary linear algebra. Our results show that under some proper assumptions, the triality theory for a class of quartic polynomial optimization problems holds strongly in the tri-duality form if the primal problem and its canonical dual have the same dimension. Otherwise, both the canonical min-max and the double-max still hold strongly, but the double-min duality holds weakly in a symmetric form.

Acknowledgments. The main results of this paper were announced in the 2nd World Congress of Global Optimization, July 3-7, 2011, Chania, Greece. The authors are sincerely indebted to Professor Hanif Sherali at Virginia Tech for his valuable comments and suggestions. David Gao's research is supported by US Air Force Office of Scientific Research under the grant AFOSR FA9550-10-1-0487. Changzhi Wu was supported by National Natural Science Foundation of China under the grant # 11001288, the Key Project of Chinese Ministry of Education under the grant # 210179, SRF for ROCS, SEM, Natural Science Foundation Project of CQ CSTC under the grant # 2009BB3057 and CMEC under the grant # KJ090802.

REFERENCES

[1] K.T. Andrews, M.F. M'Bengue and M. Shillor, *Vibrations of a nonlinear dynamic beam between two stops*, Discrete and Continuous Dynamical Systems - Series B, 12(1), (2009), 23 - 38.

[2] I. Ekeland, and R. Temam, "Convex Analysis and Variational Problems", North-Holland, 1976.

[3] S.C. Fang, D.Y. Gao, R.L. Sheu and S.Y. Wu, *Canonical dual approach for solving 0-1 quadratic programming problems*, J. Ind. and Manag. Optim. 4, (2008), 125-142.

[4] D.Y. Gao, *Post-buckling analysis and anomalous dual variational problems in nonlinear beam theory* in "Applied Mechanics in Americans Proc. of the Fifth Pan American Congress of Applied Mechanics", L.A. Godoy, L.E. Suarez (Eds.), Vol. 4. The University of Iowa, Iowa city, August (1996).

[5] D.Y. Gao, *Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory*, Applied Mechanics Reviews, 50 (11), (1997), S64-S71.

[6] D.Y. Gao, *General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics*, Meccanica 34, (1999), 169–198.

[7] D.Y. Gao, "Duality Principles in Nonconvex Systems: Theory, Methods and Applications", Kluwer Academic, Dordrecht, 2000.

[8] D.Y. Gao, *Canonical dual transformation method and generalized triality theory in nonsmooth global optimization*, J. Glob. Optim., 17(1), (2000), 127–160.

[9] D.Y. Gao, *Perfect duality theory and complete solutions to a class of global optimization problems*, Optim., 52(4–5), (2003), 467–493.

[10] D.Y. Gao, *Nonconvex semi-linear problems and canonical dual solutions*, In: Gao, D.Y., Ogden, R.W. (eds.) *Advances in Mechanics and Mathematics,* vol. II, pp. 261312. Kluwer Academic, Dordrecht (2003).

[11] D.Y. Gao, *Solutions and optimality to box constrained nonconvex minimization problems*, J. Ind. Manag. Optim. 3(2), (2007), 293–304.

[12] D.Y. Gao, *Canonical duality theory: theory, method, and applications in global optimization*, Comput. Chem., 33, (2009), 1964-1972.

[13] D.Y. Gao and R.W. Ogden, *Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation*, Quart. J. Mech. Appl. Math. 61 (4), (2008), 497-522.

[14] D.Y. Gao and N. Ruan, Solutions to quadratic minimization problems with box and integer constraints, J. Glob. Optim. 47(3), (2010), 463-484.

[15] D.Y. Gao, N. Ruan,and P.M. Pardalos, *Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization*, in *Sensors: Theory, Algorithms and Applications*, P.M. Pardalos, Y.Y. Ye, V. Boginski, and C. Commander (eds). Springer, 2010.

[16] D.Y. Gao, N. Ruan, and H. Sherali, *Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality*, J. Glob. Optim. 45(3), (2009), 473-497.

[17] D.Y. Gao and H.D. Sherali, *Canonical duality: Connection between nonconvex mechanics and global optimization*, in *Advances in Appl. Mathematics and Global Optimization*, 249-316, Springer (2009).

[18] D.Y. Gao, H.D. Sherali, and G. Strang, *Canonical duality: Objectivity, Triality, and Gap functions*, in preparation.

[19] D.Y. Gao and G. Strang, *Geometric nonlinearity: Potential energy, complementary energy, and the gap function*, Quart. Appl. Math., 47(3), (1989), 487–504.

[20] D.Y. Gao, and H.F. Yu, *Multi-scale modelling and canonical dual finite element method in phase transitions of solids*, International Journal of Solids and Structures, 45, (2008), 3660–3673.

[21] J. Gallier, *The Schur complement and symmetric positive semidefinite (and definite) matrices*, www.cis.upenn.edu/˜ jean/schurcomp.pdf.

[22] R.A. Horn, and C.R. Johnson, "Matrix Analysis", Cambridge University Press, 1985.

[23] A. Jaffe, *Constructive quantum field theory*, Mathematical Physics 2000, 111-127, Edited by A. Fokas, A. Grigoryan, T. Kibble, and B. Zegarlinski, Imperial College Press, London 2000.

[24] T.W.B. Kibble, *Phase transitions and topological defects in the early universe*, Aust. J. Phys., 50, (1997), 697-722.

[25] S.F. Li and A. Gupta, *On dual configuration forces*, J. of Elasticity, 84, (2006),13-31.

[26] J.S. Rowlinson, *Translation of J.D. van der Waals' "The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density"*, J. Statist. Phys., 20, (1979), 197-244.

[27] N. Ruan, D.Y. Gao, and Y. Jiao *Canonical dual least square method for solving general nonlinear systems of equations*, Comput. Optim. Appl. 47 (2), (2010), 335-347.

[28] D.M.M. Silva, and D.Y. Gao, *Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in* $\mathbb{R}^n$, to appear in J. Math. Analy. Appl.

[29] G. Strang, G.: Introduction to Applied Mathematics, Wellesley-Cambridge Press, 1986, 758 pp.

[30] R. Strugariu, M.D. Voisei, and C. Zalinescu, *Counter-examples in bi-duality, triality and tri-duality*, http://www.math.uaic.ro/ zalinesc/papers3.php?file=svz.pdf

[31] M.D. Voisei, and C. Zalinescu, *Some remarks concerning Gao-Strang's complementary gap function*, Applicable Analysis, (2010). DOI: 10.1080/00036811.2010.483427.

[32] M.D. Voisei, and C. Zalinescu, *Counterexamples to some triality tri-duality results*, J. Glob. Optim., DOI 10.1007/s10898-010-9592-y.

[33] Z.B. Wang, S.C. Fang, D.Y. Gao,and W.X. Xing, *Canonical dual approach to solving the maximum cut problem*, to appear in *J. Glob. Optim.*

[34] S.T. Yau, and D.Y. Gao, *Obstacle problem for von Karman equations*, Adv. Appl. Math., 13, (1992), 123-141.

E-mail address: d.gao@ballarat.edu.au
E-mail address: changzhiwu@yahoo.com